PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Repeated Cross-Sectional Surveys of Sunbed Use 2007-15 and Skin Cancer Projections of Campaign Results 2007-40 in the Danish Population
AUTHORS	Køster, Brian; Meyer, Maria; Andersson, Therese; Engholm, Gerda; Dalum, Peter

VERSION 1 – REVIEW

REVIEWER	Dr. Rüdiger Greinert Dept. Mol. Cellbiologie Skin Cancer center Buxthude, Elbekliniken Stade/Buxtehude Am Krankenhaus 121614 Buxtehude, Germany
REVIEW RETURNED	28-Feb-2018

GENERAL COMMENTS	In this manuscript the authors evaluate the Danish anti-sunbed campaign effects 2007-2015 on prevalence of sunbed use and use this results to model future effects on skin cancer incidence in a longitudal, cross-sectional design. They show that the 2007-2015 campaign was able to reduce the prevalence of sunbed use in Denmark to 30% of the precampaign level. Taking this results the authors estimate a reduced number of skin cancers cases (BCC+SCC*MM) with more than 10.000 totally during the years 2007 - 2014 using the "Prevent" model. The number would even increase to 40.000 totally reduced skin cancer cases during 2007-2040 if the 2007-2015 campaign would be continued. These are important data and estimates which can be used for further preventive efforts. However the used model (Prevent) is not explained well enough in the manuscript. It's not enough to cite to references ([24, 25] page 6. line 14). Further description should be given in this manuscript. Also the estimates for latency times (LAT and LAG) which are used for modelling further trends in MM-, BCC- and SCC-incidence (page 7, line 5 and following) are not defind precicely enough and the values taken then (page 7, line 14-16) have been deduce from literature values. However the reader would like to have some more information about the literature cited. These drawbacks in the maunscrip make it difficult to follow and to understand the manuscript fully (see bullet points 4, 7 and 10 in the checklist above). Therefore, these items should be changed.

REVIEWER	Olaf Gefeller Department of Medical Informatics, Biometry and Epidemiology, University of Erlangen
REVIEW RETURNED	12-Mar-2018

GENERAL COMMENTS	Report on bmjopen-2018-022094 (Development in Sunbed Use 2007-15 and Skin Cancer Projections of Campaign Results 2007-

	40 in the Danish Population) In its first part the manuscript addresses the temporal development of sunbed use in Denmark during the 9-year period from 2007-15. During the same period sun protection campaigning (focusing on sunbed (mis-)use) was steadily performed on the Danish population level. In its second part the manuscript tries to predict the future development of skin cancer incidence based on different scenarios. The topic of the manuscript is important and timely. The authors can report results from an impressively large data base of repeated cross-sectional studies which is definitely interesting to epidemiologists and public health scientists. My critical remarks below intend to improve the presentation of findings of this important study and to point to missing relevant information and slight errors in the description. First of all, I have to express my astonishment why the authors decided to combine the two topics (data analysis of the surveys and results from modelling the future development) into one manuscript. Putting this material into one manuscript leads to the situation that the degree of detail when describing methods and results for the two distinct topics gets too low due space constraints. Specific remarks: - p. 3, l. 4: I have never heard of a "longitudinal, cross-sectional design", it sounds like an oxymoron to me. I suggest instead "repeated cross-sectional design". - p. 3, l. 13: The statement "The prevalence of sunbed use in Denmark was reduced to 30 % of the pre-campaign level" needs clarification in two aspects: (i) it should be made clear whether sunbed use during the last 12 months or sunbed use during life is meant, (ii) absolute figures should be given, not only the proportional reduction. - p. 3, l. 23-29: The section should be completely rewritten as in its present form it does reflect the primary strengths and limitations of the study/manuscript. In its current wording the first two aspects focus on the strengths of the campaign.

	- p. 4, l. 2: Instead of "main risk factor" you should write main modifiable (or environmental) risk factor (as skin type and number of nevi show stronger effects on melanoma risk than UVR exposure). - p. 4, Introduction: The descriptive epidemiology on skin cancer has a "Denmark bias". It should be noted that the Denmark is quite typical for higher latitude countries. Look at Erdmann et al. (2013, Int J Cancer) to see the similarities between Denmark and other countries. - p. 4, Campaign content: It is not clear whether the first cross-sectional study in March 2007 took place prior to the campaign or not. You did not give the month when the campaign started. - p. 5, Questionnaire and confounding: Information on the response rates in the different surveys is not given and has to supplemented. It is of special interest whether the response rate has declined over time. If some information about non-responders is available, it should be integrated in the manuscript. - p. 5, l. 29-31: What about skin type V and VI? If skin type I to IV was an eligibility criterion for participating in the study, it should be mentioned. - p. 6, Analysis: You have defined two dichotomous outcome variables, 'ever sunbed use' and 'recent sunbed use', from your three categories 'recent users', 'non-recent users' and 'never users'. As a result of your definition you do not have the same reference group in your logistic regression models as in one model the reference group comprises only the 'never users' whereas in the other model it comprises the combination of 'non-recent users' and 'never users'. As a consequence the (crude and adjusted) ORs from table 2 and table S2 cannot be compared (they are drastically different for many factors, but the differences cannot be interpreted due the different definition of reference groups). What was the rationale for choosing this approach? Did you consider modeling the outcome variable 'sunbed use' with three categories in a single polytomous logistic model instead of creating two models with dichotomized outcome variable derived from the original variable? In addition, I would like to know why you decided to give more prominence to the 'ever sunbed use' variable in your manuscript (results for 'recent sunbed use' are shifted to the supplementary material). To my opinion, the 'recent sunbed use' variable would be much more appropriate to evaluate the campaign effect.

	- p.6, The prevent model: In general, I have no sympathy for modelling future developments as the results do depend heavily on the assumptions incorporated into the model. Is there a published methodologic paper on this "prevent model" giving the details? You are referencing a website where one can get the software and an earlier paper where this modelling approach has already been used. From your description I have only a vague idea how the model actually looks like. p. 6, l. 17-21: You write, "If the scenario of interest is no exposure or exposure with minimum impact on risk, this percentage is interpretable as the population attributable fraction (PAF) of sunbed use experience, respectively, on skin cancer (MM, SCC, BCC) incidence by the year 2040: they represent the numbers of cases that would be prevented had the population not used sunbed and therefore the fraction of MM, SCC and BCC cases attributable to these risk factors" The explanation is not entirely correct. The PAF considers the population level and tries to answers the question how many of the diseased cases in this population can be attributed to some exposure and are potentially preventable given the exposure is eliminated. However, the multifactorial etiology of almost all diseases in general and skin cancer in particular has to be taken into account when considering the effects of the elimination. Even when adjusted PAFs are considered the problem is not solved as the elimination of the exposure will lead to a change in the distribution of the confounding variables in practice. Therefore, the PAF calculation will not give a valid answer to the question on the "number of cases that would be prevented had the population not used sunbeds". Due to lack of detail in the description of the "prevent model" I cannot assess how this problem is dealt with methodologically. In any case, a more cautious wording in the manuscript when describing the results of the predicted numbers of cases is necessary. - p. 6, Incidence data: I am not familiar with the Danish cancer registry, but from my experience with cancer registry data I know that data on C44 are mostly unreliable due to severe underreporting of C44 cases. Is the situation in Denmark different? - p. 7/8: You should be always specific whether you address 'recent sunbed use' or 'ever sunbed use'. For example, your finding that "In 2015, the level of sunbed use had approximately decreased to 30 % of the pre-campaign measurement in March 2007" refers to 'recent sunbed use' which can only be realized when looking at Figure 2a/b. As already mentioned before you should also give the absolute

	figures, not only proportional reductions. BTW: Methodologically, an OR compares the odds of sunbed use and not the prevalence of sunbed use. As sunbed use is not rare in the population, there is a non-negligible numerical difference between an OR and a PR (prevalence ratio). - p. 8, l. 23-25: The impact of weather parameters on sunbed use, especially its direction, surprised me. I had expected that the tendency to use sunbeds would be lower during a warm period with many hours of sunshine, but your data tell a different story. Has a similar association been observed in other studies? - p. 8, l. 29/30: This is not a sentence! - p. 8, l. 30/31: As already mentioned above, an OR of 0.94 (or 0.82) cannot directly be interpreted as showing an annual reduction of 6% (or 18%) in the level of sunbed use. - p. 9, l. 5: Again, you report your 30% reduction omitting the absolute figures and without clearly stating that it refers to sunbed use during the last 12 months. - p. 9/10, Discussion: You did not discuss your findings related to the subgroups which reduced their sunbed use most. Are your findings in line with other studies? - Reference 31: missing volume and missing page number - Table 2/S2, right part: In the text you refer to the results reported here as OR, but in the tables you label them as 'annual percentage decrease'. In the text you interpret a smaller number as showing a stronger reduction than a higher number, which corresponds to an OR. Then your table has to be revised. Overall: There are several typos in the manuscript. Proof-reading has to be intensified.

VERSION 1 – AUTHOR RESPONSE

Overall, there was a mistake in the manuscript as pointed out by reviewer 2, which has led to corrections in e.g. table 2 and figure3. This means that the overall results from the projections are scaled down, however conclusions remain the same.

Rüdiger Greinert

In this manuscript the authors evaluate the Danish anti-sunbed campaign effects 2007-2015 on prevalence of sunbed use and use this results to model future effects on skin cancer incidence in a longitudal, cross-sectional design.

They show that the 2007-2015 campaign was able to reduce the prevalence of sunbed use in Denmark to 30% of the precampaign level. Taking this results the authors estimate a reduced number of skin cancers cases (BCC+SCC*MM) with more than 10.000 totally during the years 2007 - 2014 using the "Prevent" model. The number would even increase to 40.000 totally reduced skin cancer cases during 2007-2040 if the 2007-2015 campaign would be continued.

These are important data and estimates which can be used for further preventive efforts. However the used model (Prevent) is not explained well enough in the manuscript. It's not enough to cite to references ([24, 25] page 6. line 14). Further description should be given in this manuscript.

The model is combining population development, cancer trends, and (change in) distribution of risk factors. We have only added a little more information about the model, because the model has been already applied in dozens of studies to date and that reference 24 is a very thorough model paper as how to use Prevent including calculations. It will not be possible to have such thorough description in each application of the model, I believe. Instead, we have provided a more cautious interpretation in the discussion because as reviewer 2 points out it may lack precision due to unknown unpredictable factors. We have added the appendix2 from the model paper referenced in the end of this response.

Also the estimates for latency times (LAT and LAG) which are used for modelling further trends in MM-, BCC- and SCC-incidence (page 7, line 5 and following) are not defind precicely enough and the values taken then (page 7, line 14-16) have been deduce from literature values. However, the reader would like to have some more information about the literature cited.

The time from exposure of sunbed use to development of skin cancer is not known precisely. Neither is the time from ending exposure of UVR risk to a risk reduction. The majority of studies have used binary outcomes ever/never use. The LAT and LAG times for skin cancers are as stated not determined for a specific time in the literature and therefore these are also subject for the sensitivity analysis as shown in table 3 and as stated in last paragraph of method section. The lack of knowledge about these were added to limitations. The important thing I think is how varying the ‘LAT+LAG’ variable influences results which is shown in table 3. For e.g. BCC the fraction avoided is 0.5%, 0.6% and 0.8% for LAT+LAG times of respectively 0, 10 and 20 years. We have added about the cited studies.

These drawbacks in the maunscript make it difficult to follow and to understand the manuscript fully (see bullet points 4, 7 and 10 in the checklist above). Therefore, these items should be changed.

Study design 4 Present key elements of study design early in the paper P5L6 I believe we have provided this point. I have answered regarding response rates under reviewer 2 if this is among the requests?

Variables 7 Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable P5L12-P5L36

I believe everything which is required here is present including all outcomes, exposures, predictors, confounders and effect modifiers.

Study size 10 Explain how the study size was arrived at As described in the checklist the total study size was described in P5L12-23 and number of participants in annual data collection are shown in table1. I think we have provided this point.

Olaf Gefeller

In its first part the manuscript addresses the temporal development of sunbed use in Denmark during the 9-year period from 2007-15. During the same period sun protection campaigning (focusing on sunbed (mis-)use) was steadily performed on the Danish population level. In its second part the manuscript tries to predict the future development of skin cancer incidence based on different scenarios.
The topic of the manuscript is important and timely. The authors can report results from an impressively large data base of repeated cross-sectional studies which is definitely interesting to epidemiologists and public health scientists.
My critical remarks below intend to improve the presentation of findings of this important study and to point to missing relevant information and slight errors in the description.
First of all, I have to express my astonishment why the authors decided to combine the two topics (data analysis of the surveys and results from modelling the future development) into one manuscript. Putting this material into one manuscript leads to the situation that the degree of detail when describing methods and results for the two distinct topics gets too low due space constraints.
Specific remarks:
- p. 3, l. 4: I have never heard of a "longitudinal, cross-sectional design", it sounds like an oxymoron to me. I suggest instead "repeated cross-sectional design". Changed to ‘repeated’
- p. 3, l. 13: The statement "The prevalence of sunbed use in Denmark was reduced to 30 % of the pre-campaign level" needs clarification in two aspects: (i) it should be made clear whether sunbed use during the last 12 months or sunbed use during life is meant, (ii) absolute figures should be given, not only the proportional reduction. Changed to absolute numbers and clarified recent sunbed use
- p. 3, l. 23-29: The section should be completely rewritten as in its present form it does reflect the primary strengths and limitations of the study/manuscript. In its current wording the first two aspects focus on the strengths of the campaign. Changed
- p. 4, l. 2: Instead of "main risk factor" you should write main modifiable (or environmental) risk factor (as skin type and number of nevi show stronger effects on melanoma risk than UVR exposure). Added Modifiable
- p. 4, Introduction: The descriptive epidemiology on skin cancer has a "Denmark bias". It should be noted that the Denmark is quite typical for higher latitude countries. Look at Erdmann et al. (2013, Int J Cancer) to see the similarities between Denmark and other countries. Added sentence based on Erdmann
- p. 4, Campaign content: It is not clear whether the first cross-sectional study in March 2007 took place prior to the campaign or not. You did not give the month when the campaign started. Added ‘May’
- p. 5, Questionnaire and confounding: Information on the response rates in the different surveys is not given and has to supplemented. It is of special interest whether the response rate has declined over time. If some information about non-responders is available, it should be integrated in the manuscript. We have added a sentence, however data is not collected the same way in webpanels as in traditional paper surveys and the traditional quality measure, response rates, is more difficult to evaluate here e.g. data are collected in dynamic quotas, which means that when a sufficient number of one quota is filled for example women aged 40-49 from capital area, this group is no longer invited or valid responses are discarded by the system. This means that there is not one responserate but differential repsonserates for maybe 40 quotas. In addition, different data providers have different ways of calculating a responserate or succesrate. However, detailed datasampling descriptions exists for every year in available reports. The range of the average responserates are about 20-30 % and we cannot rule out that there could have been declining responserates, but at the same time it is very difficult to evaluate because the total number of surveys conducted by the data providers have increased concurrently and what this means to the entire survey area we can only guess.
- p. 5, l. 29-31: What about skin type V and VI? If skin type I to IV was an eligibility criterion for participating in the study, it should be mentioned. The darker skintypes are seldom in Denmark and

as categorization is done by questionnaires, skintype IV-VI was collapsed in to a single category ‘skin type IV’ however, we have changed IV to IV-VI now.
- p. 6, Analysis: You have defined two dichotomous outcome variables, 'ever sunbed use' and 'recent sunbed use', from your three categories 'recent users', 'non-recent users' and 'never users'. As a result of your definition you do not have the same reference group in your logistic regression models as in one model the reference group comprises only the 'never users' whereas in the other model it comprises the combination of 'non-recent users' and 'never users'. As a consequence the (crude and adjusted) ORs from table 2 and table S2 cannot be compared (they are drastically different for many factors, but the differences cannot be interpreted due the different definition of reference groups). What was the rationale for choosing this approach? Did you consider modeling the outcome variable 'sunbed use' with three categories in a single polytomous logistic model instead of creating two models with dichotomized outcome variable derived from the original variable? In addition, I would like to know why you decided to give more prominence to the 'ever sunbed use' variable in your manuscript (results for 'recent sunbed use' are shifted to the supplementary material). To my opinion, the 'recent sunbed use' variable would be much more appropriate to evaluate the campaign effect. I suppose we could have analyzed data in several ways. It is correct that from a campaign perspective recent sunbed use seems like the more interesting variable, however we have some very clear results. Second, we have previously published results of recent use from the first 3 years of the campaign. Third and most important, ever-use is the most reliable variable, the variable needed for the modelling and the variable on which the cancer risk is evaluated. Recent users which stop using may start to use again after a year or in different periods of life e.g. when being single or likewise.
- p.6, The prevent model: In general, I have no sympathy for modelling future developments as the results do depend heavily on the assumptions incorporated into the model. Is there a published methodologic paper on this "prevent model" giving the details? You are referencing a website where one can get the software and an earlier paper where this modelling approach has already been used. From your description I have only a vague idea how the model actually looks like. We have stated in the paper, a model is only as good as the input data, however working with prevention it is a useful tool to visualize and sometimes necessary to estimate consequences. As mentioned for reviewer 1 reference 24 is a very thorough model paper as how to use Prevent including calculations also added in the end of this response. Additionally we have added to limitations.
- p. 6, l. 17-21: You write, "If the scenario of interest is no exposure or exposure with minimum impact on risk, this percentage is interpretable as the population attributable fraction (PAF) of sunbed use experience, respectively, on skin cancer (MM, SCC, BCC) incidence by the year 2040: they represent the numbers of cases that would be prevented had the population not used sunbed and therefore the fraction of MM, SCC and BCC cases attributable to these risk factors" The explanation is not entirely correct. The PAF considers the population level and tries to answers the question how many of the diseased cases in this population can be attributed to some exposure and are potentially preventable given the exposure is eliminated. However, the multifactorial etiology of almost all diseases in general and skin cancer in particular has to be taken into account when considering the effects of the elimination. Even when adjusted PAFs are considered the problem is not solved as the elimination of the exposure will lead to a change in the distribution of the confounding variables in practice. Therefore, the PAF calculation will not give a valid answer to the question on the "number of cases that would be prevented had the population not used sunbeds". Due to lack of detail in the description of the "prevent model" I cannot assess how this problem is dealt with methodologically. In any case, a more cautious wording in the manuscript when describing the results of the predicted numbers of cases is necessary. I agree with your considerations that there are limitations to the model and to use cautious wording in the manuscript. However, I do not agree that it will not give a valid answer and confounding variables are shown with no large changes as well as sensitivity analysis is applied. It is not an exact answer. It is an indication. We have added this as mentioned above under limitations
- p. 6, Incidence data: I am not familiar with the Danish cancer registry, but from my experience with cancer registry data I know that data on C44 are mostly unreliable due to severe underreporting of C44 cases. Is the situation in Denmark different? We have added a sentence about this. The registration of non-melanoma skin cancer C44 is probably more complete in Denmark than in most other countries. Since 2004 the cancer registration has been made by a linkage between the national hospital register, the pathology register, and the cause of death register. For both melanoma skin cancer, C43, and C44, non-melanoma skin cancer, divided into basal cell carcinoma and other non-melanoma skin cancers, mainly planocellular, registrations are also included based on a registration in the pathology register alone from 2004 and on.
- p. 7/8: You should be always specific whether you address 'recent sunbed use' or 'ever sunbed use'. For example, your finding that "In 2015, the level of sunbed use had approximately decreased to 30 %

of the pre-campaign measurement in March 2007" refers to 'recent sunbed use' which can only be realized when looking at Figure 2a/b. As already mentioned before you should also give the absolute figures, not only proportional reductions. BTW: Methodologically, an OR compares the odds of sunbed use and not the prevalence of sunbed use. As sunbed use is not rare in the population, there is a non-negligible numerical difference between an OR and a PR (prevalence ratio). We have added recent or ever to sunbed use wherever relevant. There was confusion of wording in the manuscript between OR and percentage reduction leading to a very unfortunate mistake in table 2 and figure3, which has now been corrected. Thank you for pointing that out.
- p. 8, l. 23-25: The impact of weather parameters on sunbed use, especially its direction, surprised me. I had expected that the tendency to use sunbeds would be lower during a warm period with many hours of sunshine, but your data tell a different story. Has a similar association been observed in other studies? We don't know if this is a spurius finding, but it could be caused by the myth that says people should pre-tan. If we have a bad summer in Denmark perhaps people do not need to display their pale skin in outdoor situations and thus do not pretan? I'm not aware of other studies, but I did not expect this result as well.
- p. 8, l. 29/30: This is not a sentence! Changed
- p. 8, l. 30/31: As already mentioned above, an OR of 0.94 (or 0.82) cannot directly be interpreted as showing an annual reduction of 6% (or 18%) in the level of sunbed use. Changed - see above
- p. 9, l. 5: Again, you report your 30% reduction omitting the absolute figures and without clearly stating that it refers to sunbed use during the last 12 months. Changed
- p. 9/10, Discussion: You did not discuss your findings related to the subgroups which reduced their sunbed use most. Are your findings in line with other studies? A comment was added, but as already mentioned I'm not sure how comparable the different countries/campaigns are.
- Reference 31: missing volume and missing page number added
- Table 2/S2, right part: In the text you refer to the results reported here as OR, but in the tables you label them as 'annual percentage decrease'. In the text you interpret a smaller number as showing a stronger reduction than a higher number, which corresponds to an OR. Then your table has to be revised. Changed - see above
Overall: There are several typos in the manuscript. Proof-reading has to be intensified. The manuscript was proof read

Appendix 2.

Mathematical calculation in Prevent The calculation in Prevent uses the risk factor prevalence and relative risk to calculate the trend impact fraction (TIF) and potential impact fraction (PIF). The TIF dealswith the increased or decreased incident number of cases at a certain time, due to an autonomous change in risk factor prevalence, as a proportion of the incident cases that would have occurred at that time in the absence of the (autonomous) change (or trend). The PIF deals with the incident number of cases prevented at a certain time, by an intervention to reduce risk factor prevalence, as a proportion of the incident cases that would have occurred at that time in the absence of the intervention. In Prevent, the variables in TIF and PIF are dependent on age, sex, risk factor and disease, and time. For a categorical risk factor the formula for TIF is as follows:

$$\mathrm{TIF}_t^{r,d,s,a} = \frac{\sum_{c=1}^{n} p_c RR_{c,0}^{r,j=0,d,s,a} - \sum_{c=1}^{n} p_c^* RR_{c,t}^{r,j=0,d,s,a}}{\sum_{c=1}^{n} p_c RR_{c,0}^{r,j=0,d,s,a}}$$

For a categorical risk factor the formula for PIF is as follows:

$$\mathrm{PIF}_t^{r,d,s,a} = \frac{\sum_{c=1}^{n} p_c \mathrm{RR}_{c,t}^{r,j=0,d,s,a} - \sum_{c=1}^{n} p_c^* \mathrm{RR}_{c,t}^{r,j=1,d,s,a}}{\sum_{c=1}^{n} p_c \mathrm{RR}_t^{r,j=0,d,s,a}}$$

where pc is the proportion of the risk factor in category c, RRc is the relative risk for that category, and pc* is the proportion in category c after the intervention or in the case of TIF pc* is the proportion in category c after applying an autonomous change in risk factor prevalence. r, d, s, a are indicators for risk factor, disease, sex and age, respectively, whereas j is for reference if 0 and intervention population if 1. The TIF is multiplied by the population disease measure to calculate the burden of disease for the reference population. To calculate the incidence of disease among the intervention population both TIF and PIF are applied to the specific disease estimate. The difference between the two populations represents the net effect of an intervention. The model assumes independence between risk factors, using a multiplicative model.

$$I_t^{ref} = I_0 \left(1 - \prod_r (1 - TIF_r)\right)$$

$$I_t^{int} I_0 \left(1 - \prod_r (1 - TIF_r)(1 - PIF_r)\right)$$

where I0 is the disease incidence rate in the base year and r is an index for risk factor. I ref is the incidence rate in the reference population and I int is the incidence rate in the intervention population. All measures are age- and sex-specific, and recalculated for each period year based on period-specific risk factor prevalence and constant relative risks or risk functions. The incidence rates are then applied to the population number (age-, sex-, and period-specific) to calculate the absolute changes in incidence cases due to autonomous trend or intervention.

VERSION 2 – REVIEW

REVIEWER	Dr. Rüdiger Greinert Dept. Mol. Cell Biology Skin Cancer center Buxtheude Elbekliniken Stade/Buxtehude Am Krankenhaus 1 21614 Buxtehude, Germany
REVIEW RETURNED	23-Apr-2018

GENERAL COMMENTS	None

REVIEWER	Olaf Gefeller Department of Medical Informatics, Biometry and Epidemiology, University of Erlangen
REVIEW RETURNED	18-May-2018

GENERAL COMMENTS	Report on bmjopen-2018-022094.R1 (Development in Sunbed Use 2007-15 and Skin Cancer Projections of Campaign Results 2007-40 in the Danish Population)

	In the revised version of their manuscript the authors have addressed most of the reviewers' remarks satisfactorily. I am pleased with the explanations added to the manuscript and the correction of the confusion between OR and annual percentage reductions in Tab. 2/Fig. 3. In addition to some minor points (see below), only one major point still remains. Major: As already criticized in the first review the authors use different reference groups for the dichotomized outcome variables related to sunbed use in their logistic regression models. As a consequence the effect estimates (crude and adjusted ORs) of explanatory variables cannot be compared between the models for recent sunbed use and ever sunbed use (Tab. 2 and Tab. S2). The authors' reply addressing my question about the rationale for this unusual methodologic approach („I suppose we could have analyzed the data in several ways") is not really enlightening. If the authors have no good reason for their approach, then they should reanalyze their data in a way offering comparability of effect estimates between models for recent and ever sunbed use (e.g. by eliminating non-recent sunbed users from the reference group in the logistic model for recent sunbed use or by using a single polytomous logistic model for an outcome variable with three outcome categories, namely 'recent use', 'non-recent use' and 'never use'). Minor: - In the corrected version the estimated ORs for the variable 'sunbed use annual change' show – at first glance – a very limited numerical variability in the subgroups. This impression is induced by the choice of the short time period of one year. The definition of a longer time period for such a variable (e.g. 'change in sunbed use over five years') would lead to numerically lower OR estimates and a more heterogeneous pattern of effect estimates in subgroups highlighting existing differences. - p. 16: Figure legend for Figure 2A and 2B are identical.

	- p. 19: The graphical displays have headlines stating that both figures (left and right) are Figure 2A.

VERSION 2 – AUTHOR RESPONSE

Report on bmjopen-2018-022094.R1 (Development in Sunbed Use 2007-15 and Skin Cancer Projections of Campaign Results 2007-40 in the Danish Population)
In the revised version of their manuscript the authors have addressed most of the reviewers' remarks satisfactorily. I am pleased with the explanations added to the manuscript and the correction of the confusion between OR and annual percentage reductions in Tab. 2/Fig. 3. In addition to some minor points (see below), only one major point still remains.
Thank you for the comments.
Major:
As already criticized in the first review the authors use different reference groups for the dichotomized outcome variables related to sunbed use in their logistic regression models. As a consequence the effect estimates (crude and adjusted ORs) of explanatory variables cannot be compared between the models for recent sunbed use and ever sunbed use (Tab. 2 and Tab. S2). The authors' reply addressing my question about the rationale for this unusual methodologic approach („I suppose we could have analyzed the data in several ways") is not really enlightening. If the authors have no good reason for their approach, then they should reanalyze their data in a way offering comparability of effect estimates between models for recent and ever sunbed use (e.g. by eliminating non-recent sunbed users from the reference group in the logistic model for recent sunbed use or by using a single polytomous logistic model for an outcome variable with three outcome categories, namely 'recent use', 'non-recent use' and 'never use').
Sorry, the answer was inappropriate. What we meant was, that these data have many possibilities and your suggestion is also very interesting and would provide valuable information. This paper, however, addresses the development in sunbed use in the population, both recent and ever sunbed use. It is true the OR's from the 2 models are not comparable. However, our main aim is not to examine determinants for sunbed use, nor to compare the models. We are establishing models for the development of sunbed use in the population during 2007-15. Leaving out part of the population would influence the model and the final result. Even though 'ever use' is the main input for the projection model, we are also interested in the development of 'recent sunbed use'. Therefore it is essential to keep the model as it is.

Minor:
- In the corrected version the estimated ORs for the variable 'sunbed use annual change' show – at first glance – a very limited numerical variability in the subgroups. This impression is induced by the choice of the short time period of one year. The definition of a longer time period for such a variable (e.g. 'change in sunbed use over five years') would lead to numerically lower OR estimates and a more heterogeneous pattern of effect estimates in subgroups highlighting existing differences.
Agree, however using 5 years could also be interpreted as exaggerating the effects. We think the one year intervals is appropriate. It is used for the projections and focus is kept on the 'reel' differences, e.g. young age group.
- p. 16: Figure legend for Figure 2A and 2B are identical.
Ever/Recent added
- p. 19: The graphical displays have headlines stating that both figures (left and right) are Figure 2A.

Changed